Multivariate Normal Distribution I

Nipun Batra

IIT Gandhinagar

July 31, 2025

Outline

Univariate Normal Distribution

Univariate Gaussian PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Univariate Normal Distribution

Univariate Gaussian PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Univariate Normal Distribution

Univariate Gaussian PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Notation: $X \sim \mathcal{N}(\mu, \sigma^2)$

Univariate Normal Distribution

Univariate Gaussian PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Notation: $X \sim \mathcal{N}(\mu, \sigma^2)$
Parameters:

- $\mu \in \mathbb{R}$: **Mean** (location parameter)

Univariate Normal Distribution

Univariate Gaussian PDF

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Notation: $X \sim \mathcal{N}(\mu, \sigma^2)$
Parameters:

- $\mu \in \mathbb{R}$: **Mean** (location parameter)
- $\sigma^2 > 0$: **Variance** (spread parameter)

Pop Quiz: The Normalizing Constant

Quick Quiz 1

Why does the Gaussian PDF have $\frac{1}{\sigma\sqrt{2\pi}}$ in the denominator?

a) It makes the formula look nice

Answer: b) It's the normalizing constant that makes $\int_{-\infty}^{\infty} f(x)dx = 1$!

Pop Quiz: The Normalizing Constant

Quick Quiz 1

Why does the Gaussian PDF have $\frac{1}{\sigma\sqrt{2\pi}}$ in the denominator?

a) It makes the formula look nice
b) It ensures the total probability integrates to 1

Answer: b) It's the normalizing constant that makes $\int_{-\infty}^{\infty} f(x)dx = 1$!

Pop Quiz: The Normalizing Constant

Quick Quiz 1

Why does the Gaussian PDF have $\frac{1}{\sigma\sqrt{2\pi}}$ in the denominator?

a) It makes the formula look nice
b) It ensures the total probability integrates to 1
c) It's related to the standard deviation

Answer: b) It's the normalizing constant that makes $\int_{-\infty}^{\infty} f(x)dx = 1$!

Why $\frac{1}{\sigma\sqrt{2\pi}}$? Let's Derive It!

Setup: Let normalizing constant be c and $g(x) = e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$1 = \int_{-\infty}^{\infty} c \cdot g(x) dx$$

Why $\frac{1}{\sigma\sqrt{2\pi}}$? Let's Derive It!

Setup: Let normalizing constant be c and $g(x) = e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$1 = \int_{-\infty}^{\infty} c \cdot g(x) dx$$
$$1 = \int_{-\infty}^{\infty} c e^{-\frac{(x-u)^2}{2\sigma^2}} dx$$

Why $\frac{1}{\sigma\sqrt{2\pi}}$? Let's Derive It!

Setup: Let normalizing constant be c and $g(x) = e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$1 = \int_{-\infty}^{\infty} c \cdot g(x) dx$$

$$1 = \int_{-\infty}^{\infty} c e^{-\frac{(x-u)^2}{2\sigma^2}} dx$$

Let's substitute $\frac{x-u}{\sqrt{2}\sigma}$ with t.

Why $\frac{1}{\sigma\sqrt{2\pi}}$? Let's Derive It!

Setup: Let normalizing constant be c and $g(x) = e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$1 = \int_{-\infty}^{\infty} c \cdot g(x) dx$$

$$1 = \int_{-\infty}^{\infty} c e^{-\frac{(x-u)^2}{2\sigma^2}} dx$$

Let's substitute $\frac{x-u}{\sqrt{2}\sigma}$ with t.

$$1 = \int_{-\infty}^{\infty} c e^{-t^2} dt \times \sqrt{2}\sigma$$

Why $\frac{1}{\sigma\sqrt{2\pi}}$? Let's Derive It!

Setup: Let normalizing constant be c and $g(x) = e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$$1 = \int_{-\infty}^{\infty} c \cdot g(x)dx$$

$$1 = \int_{-\infty}^{\infty} ce^{-\frac{(x-u)^2}{2\sigma^2}}dx$$

Let's substitute $\frac{x-u}{\sqrt{2}\sigma}$ with t.

$$1 = \int_{-\infty}^{\infty} ce^{-t^2}dt \times \sqrt{2}\sigma$$

$$1 = \sqrt{2}\sigma c \times 2\int_{0}^{\infty} e^{-t^2}dt$$

Univariate Normal Distribution

$$\frac{2}{\sqrt{\pi}} \int_{0}^{\infty} e^{-t^2} dt$$

The above expression is called error function and is it's value is denoted by $erf(t)$. In our case, we want $erf(\infty)$ which is equal to 1.

Univariate Normal Distribution

$$\frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-t^2} dt$$

The above expression is called error function and is it's value is denoted by $erf(t)$. In our case, we want $erf(\infty)$ which is equal to 1.

$$1 = \sqrt{2\pi}\sigma c \times \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-t^2} dt$$

Univariate Normal Distribution

$$\frac{2}{\sqrt{\pi}} \int_0^\infty e^{-t^2} dt$$

The above expression is called error function and is it's value is denoted by $erf(t)$. In our case, we want $erf(\infty)$ which is equal to 1.

$$1 = \sqrt{2\pi}\sigma c \times \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-t^2} dt$$
$$1 = \sqrt{2\pi}\sigma c \times 1$$

Univariate Normal Distribution

$$\frac{2}{\sqrt{\pi}} \int_0^\infty e^{-t^2} dt$$

The above expression is called error function and is it's value is denoted by $erf(t)$. In our case, we want $erf(\infty)$ which is equal to 1.

$$1 = \sqrt{2\pi}\sigma c \times \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-t^2} dt$$

$$1 = \sqrt{2\pi}\sigma c \times 1$$

$$\frac{1}{\sqrt{2\pi}\sigma} = c$$

Univariate Normal Distribution

Univariate Normal Distribution

Univariate Normal Distribution

Bivariate Normal Distribution

Bivariate normal distribution of two-dimensional random vector $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}_{\mathbf{2}}(\mu, \ )$$

where, mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_2] \end{bmatrix}$
and, covariance matrix Σ

$$\Sigma_{i,j} := \mathrm{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathrm{Cov}[X_i, X_j]$$

Bivariate Normal Distribution

Question: What is Cov(X, X)?

Bivariate Normal Distribution

Question: What is Cov(X, X)?
Answer: Var(X) = Cov(X, X) = $\mathrm{E}[(X - \mathrm{E}[X])]^2$
In the case of univariate normal, Var(X) is written as σ^2

Bivariate Normal Distribution

Question: What is Cov(X, X)?
Answer: Var(X) = Cov(X, X) = $\mathrm{E}[(X - \mathrm{E}[X])]^2$
In the case of univariate normal, Var(X) is written as σ^2
Question: What is the relation between $\Sigma_{i,j}$ and $\Sigma_{j,i}$?

Bivariate Normal Distribution

Question: What is Cov(X, X)?
Answer: Var(X) = Cov(X, X) = $\mathrm{E}[(X - \mathrm{E}[X])]^2$
In the case of univariate normal, Var(X) is written as σ^2
Question: What is the relation between $\Sigma_{i,j}$ and $\Sigma_{j,i}$?
Answer: They are the same!

Bivariate Normal Distribution

Question: What is Cov(X, X)?
Answer: Var(X) = Cov(X, X) = $\mathrm{E}[(X - \mathrm{E}[X])]^2$
In the case of univariate normal, Var(X) is written as σ^2
Question: What is the relation between $\Sigma_{i,j}$ and $\Sigma_{j,i}$?
Answer: They are the same!
Question: What can we say about the covariance matrix Σ?

Bivariate Normal Distribution

Question: What is Cov(X, X)?
Answer: Var(X) = Cov(X, X) = $\mathrm{E}[(X - \mathrm{E}[X])]^2$
In the case of univariate normal, Var(X) is written as σ^2
Question: What is the relation between $\Sigma_{i,j}$ and $\Sigma_{j,i}$?
Answer: They are the same!
Question: What can we say about the covariance matrix Σ?
Answer: It is symmetric. Thus $\Sigma = \Sigma^T$

Correlation and Covariance

If X and Y are two random variables, with means (expected values) μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively, then their covariance and correlation are as follows:

$$\text{cov}_{XY} = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

Correlation and Covariance

If X and Y are two random variables, with means (expected values) μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively, then their covariance and correlation are as follows:

$$\text{cov}_{XY} = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

$$\text{corr}_{XY} = \rho_{XY} = E[(X - \mu_X)(Y - \mu_Y)]/(\sigma_X \sigma_Y)$$

Correlation and Covariance

If X and Y are two random variables, with means (expected values) μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively, then their covariance and correlation are as follows:

$$\text{cov}_{XY} = \sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)]$$

$$\text{corr}_{XY} = \rho_{XY} = E[(X - \mu_X)(Y - \mu_Y)]/(\sigma_X \sigma_Y)$$

so that

$$\rho_{XY} = \sigma_{XY}/(\sigma_X \sigma_Y)$$

where E is the expected value operator.

PDF of bivariate normal distribution

We might have seen that

$$f_X(X_1, X_2) = \frac{exp(\frac{-1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu))}{2\pi|\Sigma|^{\frac{1}{2}}}$$

How do we get such a weird looking formula?!

PDF of bivariate normal with no cross-correlation

Let us assume no correlation between X_1 and X_2.
We have $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$
We have $f_X(X_1, X_2) = f_X(X_1)f_X(X_2)$

$$= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X_1-\mu_1}{\sigma_1}\right)^2} \times \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X_2-\mu_2}{\sigma_2}\right)^2}$$

$$= \frac{1}{\sigma_1\sigma_2 2\pi} e^{-\frac{1}{2}\{\left(\frac{X_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{X_2-\mu_2}{\sigma_2}\right)^2\}}$$

PDF of bivariate normal with no cross-correlation

Let us consider only the exponential part for now
$Q = \left(\frac{X_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{X_2-\mu_2}{\sigma_2}\right)^2$
Question: Can you write Q in the form of vectors X and μ?

$$= \begin{bmatrix} X_1 - \mu_1 & X_2 - \mu_2 \end{bmatrix}_{1\times 2} g(\Sigma)_{2\times 2} \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{bmatrix}_{2\times 1}$$

Here $g(\Sigma)$ is a matrix function of Σ that will result in σ_1^2 like terms in the denominator; also there is no cross-terms indicating zeros in right diagonal!

$g(\Sigma) = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{bmatrix}_{2\times 2} = \frac{1}{\sigma_1^2\sigma_2^2} \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}_{2\times 2} = \frac{1}{|\Sigma|} \mathrm{adj}(\ ) = \Sigma^{-1}$

PDF of bivariate normal with no cross-correlation

Let us consider the normalizing constant part now.
$M = \frac{1}{\sigma_1 \sigma_2 2\pi} = \frac{1}{2\pi \times |\Sigma|^{\frac{1}{2}}}$

Bivariate Gaussian samples with cross-correlation $\neq 0$

Bivariate Gaussian samples with cross-correlation = 0

Intuition for Multivariate Gaussian

Let us assume no correlation between the elements of $\mathbf{X}$.
This means Σ is a diagonal matrix.

We have $\Sigma = \begin{bmatrix} \sigma_1^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sigma_n^2 \end{bmatrix}$

And,

$$p(\mathbf{x}; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)\right)$$

As seen in the case for univariate Gaussians, we can write the following for the multivariate case,
We have $f_X(X_1, \cdots, X_n) = f_X(X_1) \times \cdots \times f_X(X_n)$

Intuition for Multivariate Gaussian

Now,

$$= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} \times \cdots \times \frac{1}{\sigma_n\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_n-\mu_n}{\sigma_n}\right)^2}$$

$$= \frac{1}{\sigma_1 \cdots \sigma_n (2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{x_n-\mu_n}{\sigma_n}\right)^2\}}$$

Taking all $\sqrt{2\pi}$ together, we get $(2\pi)^{\frac{n}{2}}$.
Similarly, taking all σ together, we get $\sigma_1 \cdots \sigma_n$. Which can be written as $|\Sigma|^{\frac{1}{2}}$, given the determinant of a digonal matrix is the multiplication of its diagonal elements.

Now, let us remove the assumption of no covariance among the elements of X

Main idea: A correlated Gaussian is a rotated independent Gaussian[1]

Rotate input space using rotation matrix R.

$$p(\mathbf{x};\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mathbf{R}^T\mathbf{x} - R^T\mu)^T\Sigma^{-1}(\mathbf{R}^T\mathbf{x} - R^T\mu)\right)$$

$$p(\mathbf{x};\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T R\Sigma^{-1}R^T(\mathbf{x} - \mu)\right)$$

[1] Neil Lawrence GPSS 2016

$C = R\Sigma^{-1}R^T$

$$p(\mathbf{x}; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|C|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T C^{-1}(\mathbf{x} - \mu)\right)$$